

(11) **EP 0 903 670 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
24.03.1999 Bulletin 1999/12

(51) Int. Cl.⁶: **G06F 13/22, G06F 13/24**

(21) Application number: **97410104.0**

(22) Date of filing: **19.09.1997**

(84) Designated Contracting States:
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
 NL PT SE**

(71) Applicant:
**Hewlett-Packard Company
 Palo Alto, California 94304 (US)**

(72) Inventor: **Delorme, Alexandre
 38000 Grenoble (FR)**

(74) Representative:
**Lloyd, Richard Graham
 Intellectual Property Section,
 Legal Department,
 HEWLETT-PACKARD FRANCE,
 Etablissement de Grenoble
 38053 Grenoble Cedex 9 (FR)**

(54) **Improved peripheral device control in data processing systems**

(57) Data processing apparatus is described comprising a processor and at least one peripheral device. The processor is arranged to service the peripheral device either in an interrupt mode in which the peripheral device is serviced in response to interrupt signals generated by the peripheral device or in a timed mode in which the peripheral device is periodically polled and serviced if required. The apparatus has a dynamic switching arrangement for switching from the interrupt mode to the timed mode depending upon conditions dynamically determined within the apparatus, at least one of said conditions being that the rate at which the peripheral device generates interrupt signals exceeds a predefined or programmable threshold frequency. The rate of polling in the timed mode is less than the threshold frequency. Thus, dynamic switching between an interrupt driven mode and a timed mode is used in order to make more efficient use of shared system resources by servicing peripheral devices only periodically at times of high demand.

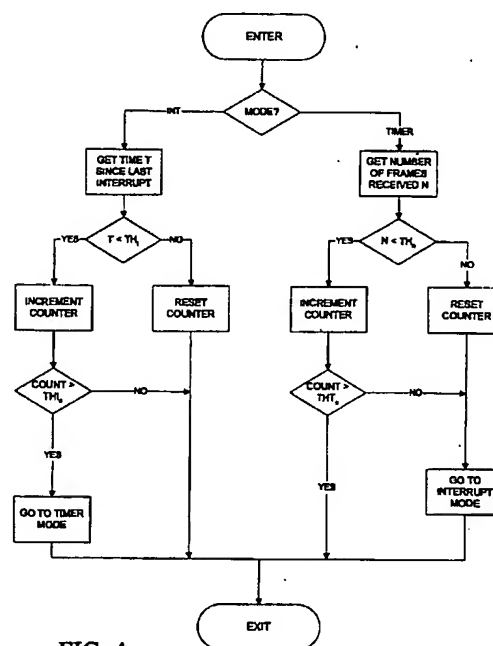


FIG. 4

EP 0 903 670 A1

1

EP 0 903 670 A1

2

Description**Field of the Invention**

[0001] The present invention relates generally to data processing apparatus comprising a processor and at least one peripheral device and, more particularly, to such apparatus in which the processor is arranged to service the peripheral device either in an interrupt mode in which the peripheral device is serviced in response to interrupt signals generated by the peripheral device or in a timed mode in which the peripheral device is periodically polled and serviced if required, the apparatus having a switching arrangement for switching from the interrupt mode to the timed mode depending upon conditions dynamically determined within the apparatus, at least one of said conditions being that the rate at which the peripheral device generates interrupt signals exceeds a predefined or programmable threshold frequency.

Background of the Invention

[0002] A data processing system of the above-described general type is proposed in US Patent number 5,414,858. In the computer system described in that patent the latency associated with interrupt processing is said to be relatively high since the interrupt signals initiate operating system routines that are relatively time consuming. On the other hand, the latency associated with polling is said to be relatively low, since the frequency of the polling can be made relatively fast.

[0003] In consequence, the patent proposes dynamically switching from an interrupt driven mode to a polling mode in order to reduce latency.

[0004] However, the power of modern processors means that the latency associated with interrupt driven processing is sufficiently low so as not to be significant for some applications. Nevertheless, even in powerful modern computer systems, interrupt processing still involves a performance overhead for the computer, for example since data associated with the task or tasks the computer is performing when it receives the interrupt need to be saved and then restored once the interrupt routine has terminated.

[0005] The problem with the approach proposed in US 5,414,858 is that, in at least some applications, the use of fast polling can also use significant processor resources and therefore the dynamic switching approach proposed would not necessarily provide any improvement in overall efficiency.

[0006] The object of this invention is to provide an improved technique for managing the servicing of peripheral devices in computer systems and which mitigates the above described drawback of the prior art.

Summary of the Invention

[0007] This invention provides a data processing system of the above described type characterised in that the rate of polling in the timed mode is less than the threshold frequency. In this way, the load on the processor caused by frequent interrupts is reduced.

[0008] Thus, dynamic switching between a interrupt driven mode and a timed mode is not used in order to reduce latency, but rather in order to make more efficient use of shared system resources by servicing peripheral devices only periodically at times of high demand.

[0009] Such an approach is particularly suitable for Input/Output (I/O) devices which exhibit a wide variation over time in the level of service they require.

[0010] In preferred embodiments, the apparatus is arranged to handle, if required, a plurality of service events in each cycle in the timed mode. In this way, the overall number of service events handled is not reduced, rather the handling of the events is scheduled so as to make more efficient use of shared system resources.

[0011] In this case, the dynamic switching arrangement can be arranged to switch from the timed mode to the interrupt mode when the number of service events handled within each polling cycle falls below a predefined or programmable threshold.

[0012] In one embodiment, the dynamic switching arrangement is arranged to measure the number of service events within each polling cycle and to switch to the interrupt mode when the number of service events handled within each polling cycle is below a predefined or programmable threshold for a predefined or programmable number of consecutive cycles.

[0013] Similarly, the dynamic switching arrangement can be arranged, when in the interrupt mode, to measure the time interval between each interrupt and to switch to the timed mode when the time interval between each interrupt is below a predefined or programmable threshold for a predefined or programmable number of consecutive received interrupts.

[0014] In one application for which the above technique is particularly suited, the peripheral device is a network adapter including a buffer memory and means to receive and store in the buffer memory frames of data received from a data communications network. In the following, the word frame will be used as a generic term to describe the units in which data is transferred across any particular network, although it will be appreciated that the proper technical term for such units may vary in different network systems. In this case, the service events referred to above are the frames of data received from the network requiring storage in a data storage device. In this implementation, the adapter is settable either to transmit or not to transmit an interrupt signal to the processor whenever a frame of data is received from the network and the switching arrangement com-

prises means to set the network adapter not to transmit the interrupt signal upon switching to the timed mode.

[0015] Viewed from another aspect, the invention also provides a method for operating data processing apparatus comprising a processor and at least one peripheral device, the method comprising servicing the peripheral device either in an interrupt mode in which the peripheral device is serviced in response to interrupt signals generated by the peripheral device or in a timed mode in which the peripheral device is periodically polled and serviced if required: and dynamically switching from the interrupt mode to the timed mode depending upon conditions dynamically determined within the apparatus, at least one of said conditions being that the rate at which the peripheral device generates interrupt signals exceeds a predefined or programmable threshold frequency, wherein the rate of polling in the timed mode is less than the threshold frequency.

[0016] Also provided is a computer program product for execution on data processing apparatus comprising a processor and at least one peripheral device to carry out the above-described method.

Brief Description of the Drawings

[0017] A data communications system embodying the invention will now be described, by way of non-limiting example, with reference to the accompanying diagrammatic drawings, in which:

- .Figure 1 is a schematic diagram showing a data processing system;
- .Figure 2 is a schematic diagram showing the network adapter in the system of Fig 1;
- .Figure 3 illustrates a software structure in the system of Fig 1;
- .Figure 4 is a flow diagram showing the operation of an interrupt routine in the system of Fig 1.

Best Mode of Carrying Out the Invention

[0018] A network adapter is an interface device that is located between a data communications network and a computer to be connected to the network. Data flows between the computer and the network in both directions, typically in units of fixed or variable size known as data packets or frames.

[0019] In general, one of the functions a network adapter has to perform is to deliver in a timely manner data frames received from the network to main memory in the computer, where they can be processed by an application program.

[0020] Conventionally, network adapters have used interrupt driven mechanisms to transfer data from the adapter to the computer. This means that when a data frame is received by the adapter and stored in its buffer, processing in the computer is interrupted and an interrupt routine performed which transfers the incoming

frame to the main storage of the computer.

[0021] Interrupt processing involves a performance overhead for the computer because, for example, data associated with the task or tasks the computer is performing when it receives the interrupt need to be saved and then restored once the interrupt routine has terminated.

[0022] If the rate at which frames are received is too high then this overhead can account for a significant proportion of the resources of the machine.

[0023] Described here is an adaptive switching technique implemented as part of a physical driver for a network adapter which makes more efficient use of system resources.

[0024] Referring to Fig 1, there is shown a data processing system comprising a central processing unit (CPU) 100 and a network adapter card 110 interconnected by a bus 120. In this implementation, bus 120 is the well known Personal Computer Interconnect (PCI) bus which is widely used in personal computers. CPU 100 is connected to bus 120 via bus interface 130. In conventional fashion, CPU 100 is connected to main system memory 140 via local processor bus 150.

[0025] Fig 2 shows in more detail the structure of network adapter card 110. Adapter 110 comprises PCI interface unit 200 and a wide area network (WAN) interface 210 which includes appropriate network line drivers and receivers for interfacing with a wide area network using standard physical layer interfaces such as RS232C, RS449, RS530, V35 or X21.

[0026] Also included in network adapter card 110 is RAM 230 which serves as a buffer memory and a Serial Communication Adapter (SCA) control unit 220 which controls the protocols used on the network lines. Serial Communication Adapter (SCA) control unit 220 includes a DMA controller (not shown) for transferring data to RAM 230. PCI interface unit 200 includes target only PCI interface logic and a memory controller for controlling RAM 230 (not shown). These units are interconnected by local bus 240. It will be recognised by those skilled in the art that network adapter cards having this general structure are widely commercially available.

[0027] Fig 3 is a schematic diagram showing the layered structure of the communication software which executes on CPU 100 in this implementation. Such a layered structure will be very familiar to those skilled in the art. Application programs represented at 300 interface directly to a protocol stack 310. Protocol stack 310 may be for example an implementation of an X25 or frame relay protocol. Protocol stack 310 interfaces to a physical driver 320. The role of the physical driver 320 is to interface directly to the hardware of the adapter card 110 indicated in Fig 3 at 330.

[0028] The general operation of the system described above will be well understood by those skilled in the art and need not be described in any more detail here.

[0029] The present invention is implemented in the

5

EP 0 903 670 A1

6

interrupt handling procedures of the data reception algorithm of physical driver 320 as will be described below.

[0030] The performance of the data reception algorithm is critical since the functionality of adapter card 110 is somewhat limited. In particular, adapter card 110 does not have the capability to directly transfer data to system memory 140. Rather, CPU 100 must read each inbound data frame from card memory 230 and transfer it to system memory 140. If adapter card 100 uses, as is conventional, a host interrupt to initiate such a transfer, then at a typical data transfer rate such as 4Mbps, and with a frame size of 128 bytes, the host cpu would be interrupted every 300 microseconds.

[0031] The reception algorithm of the physical driver is able to reduce this number of interrupts using a timer. The basic frame reception process is as follows. When adapter 110 receives a data frame it is placed in card memory 230. Adapter 110 is settable so that on receipt of a data frame it sends an interrupt signal to CPU 100 which executes an interrupt routine in physical driver software 320.

[0032] The interrupt routine of physical driver software 320 operates in two modes - an interrupt mode in which the adapter sends an interrupt signal to CPU 100 whenever it receives a data frame and a timed mode in which the adapter is set so as not to send any interrupts to the cpu, rather host interrupts are generated periodically under the control of the physical driver routine. In either case the same interrupt routine is executed by the cpu.

[0033] Fig 4 is a flow chart showing the interrupt routine. It can be seen that the switching between the two modes is governed by the timer period T and four threshold parameters. TH_i , TH_n , TH_c and THT_c .

[0034] In the timed mode, the number of data frames transferred to memory 140 within each polling cycle is measured. The driver switches to interrupt mode if the number of service events handled within each polling cycle is below a programmable threshold TH_n for a programmable number THT_c of consecutive cycles.

[0035] In the interrupt mode, the time interval between each interrupt is measured. The driver switches to the timed mode when the time interval between each interrupt is below a programmable threshold TH_i for a programmable number of consecutive received interrupts TH_c .

[0036] Of course, it is possible that in some implementations any of the four parameters TH_i , TH_n , TH_c and THT_c may be fixed.

[0037] In general, the timer period T needs to be longer than the threshold TH_i for the period between interrupts so that the overall number of interrupts is reduced. It will be understood that the threshold TH_i effectively defines a threshold frequency for the rate of frame receipt above which the driver will switch to operate in timed mode. The threshold TH_n defines a second threshold frequency, which may be the same as or different from that defined by TH_i . The algorithm described above provides allows a degree of inertia in the switch-

ing to avoid excessive and unnecessary switching between the two modes. The count thresholds THT_c and TH_c ensure that any change in the rate of receipt of frames needs to endure for a certain time before initiating a switch from one mode to the other.

[0038] The maximum latency that can be permitted for frame receipt determines an upper bound to the timer period T.

[0039] For example, timer period T may be set to 1mS so that at the highest data transfer rate of 4Mbps, 3 or 4 128-byte frames are transferred in each cycle in the timed mode. In this example, TH_i might be set to 500 microseconds and TH_c to 10, so that if more than 10 consecutive frames are received with a spacing of less than 500 microseconds, the driver switches to the timed mode. TH_n might be set to 1 and THT_c to 10, so if in more than 10 consecutive timer periods less than 1 frame is handled then the driver switches back to interrupt mode.

[0040] Of course, the parameters T, TH_i , TH_n , TH_c and THT_c may all be adjusted so as to optimise the performance according to expected network traffic conditions and or processor load. This adjustment could, for example, be carried out manually by a user as part of a system set up process or could be carried out dynamically by either an application program or the physical driver itself.

[0041] It will be appreciated that the physical driver of the present embodiment takes the form of a computer program which may be marketed in the form of a suitable computer program product including the functionality described. It will be appreciated that the invention may equally be implemented as special purpose hardware or any combination of software and hardware.

[0042] Although a specific embodiment of the invention has been described, the invention is not to be limited to the specific arrangement so described. The invention is limited only by the claims.

40 Claims

1. Data processing apparatus comprising a processor and at least one peripheral device, the processor being arranged to service the peripheral device either in an interrupt mode in which the peripheral device is serviced in response to interrupt signals generated by the peripheral device or in a timed mode in which the peripheral device is periodically polled and serviced if required, the apparatus having a dynamic switching arrangement for switching from the interrupt mode to the timed mode depending upon conditions dynamically determined within the apparatus, at least one of said conditions being that the rate at which the peripheral device generates interrupt signals exceeds a predefined or programmable threshold frequency, wherein the rate of polling in the timed mode is less than the threshold frequency.

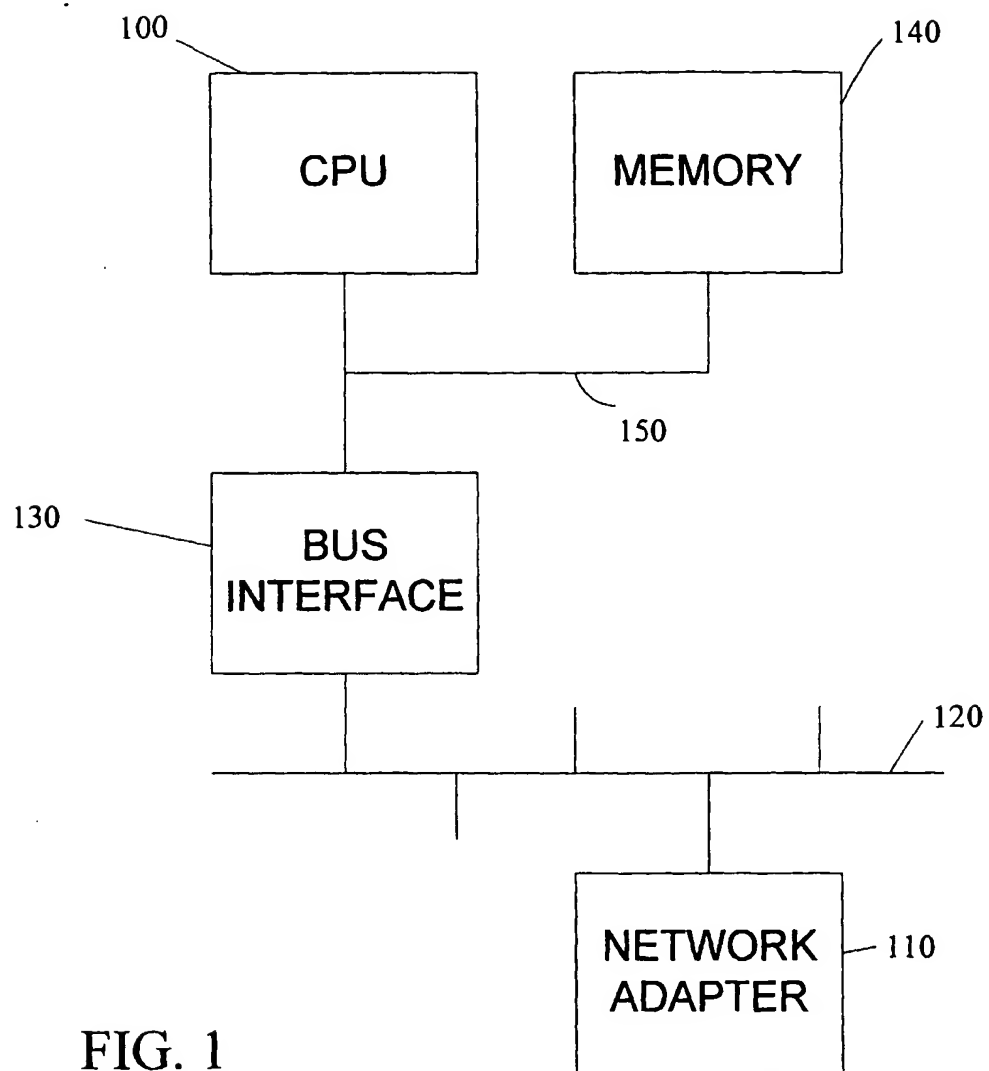
7

EP 0 903 670 A1

8

2. Data processing apparatus as claimed in claim 1 wherein, in the timed mode, the apparatus is arranged to, if required, handle a plurality of service events in each cycle.
3. Data processing apparatus as claimed in claim 2 wherein the dynamic switching arrangement is arranged to switch from the timed mode to the interrupt mode when the number of service events handled within each polling cycle falls below a predefined or programmable threshold.
4. Data processing apparatus as claimed in claim 3 wherein the dynamic switching arrangement is arranged to measure the number of service events within each polling cycle and to switch to the interrupt mode when the number of service events handled within each polling cycle is below a predefined or programmable threshold for a predefined or programmable number of consecutive cycles.
5. Data processing apparatus as claimed in claim 1 wherein the dynamic switching arrangement is arranged, when in the interrupt mode, to measure the time interval between each interrupt and to switch to the timed mode when the time interval between each interrupt is below a predefined or programmable threshold for a predefined or programmable number of consecutive received interrupts.
6. Data processing apparatus as claimed in claim 2 including a data storage device and wherein the peripheral device is a network adapter including a buffer memory and means to receive and store in the buffer memory frames of data received by the adapter from a data communications network, the service events being the frames of data received from the network requiring storage in the data storage device, wherein the adapter is settable either to transmit or not to transmit an interrupt signal to the processor whenever a frame of data is received from the network, the switching arrangement comprising means to set the network adapter not to transmit the interrupt signal upon switching to the timed mode.
7. Data processing apparatus as claimed in claim 1 comprising means for setting at least one of the polling rate and the threshold frequency according to conditions dynamically determined within the apparatus.
8. A method for operating data processing apparatus comprising a processor and at least one peripheral device, the method comprising servicing the peripheral device either in an interrupt mode in which the peripheral device is serviced in response to interrupt signals generated by the peripheral device or in a timed mode in which the peripheral device is periodically polled and serviced if required, and dynamically switching from the interrupt mode to the timed mode depending upon conditions dynamically determined within the apparatus, at least one of said conditions being that the rate at which the peripheral device generates interrupt signals exceeds a predefined or programmable threshold frequency, wherein the rate of polling in the timed mode is less than the threshold frequency.
9. A computer program product for execution on data processing apparatus comprising a processor, at least one peripheral device, and means for servicing the peripheral device either in an interrupt mode in which the peripheral device is serviced in response to interrupt signals generated by the peripheral device or in a timed mode in which the peripheral device is periodically polled and serviced if required, the program product comprising program logic for dynamically switching from the interrupt mode to the timed mode depending upon conditions dynamically determined within the apparatus, at least one of said conditions being that the rate at which the peripheral device generates interrupt signals exceeds a predefined or programmable threshold frequency, wherein the rate of polling in the timed mode is less than the threshold frequency.

EP 0 903 670 A1

**FIG. 1**

EP 0 903 670 A1

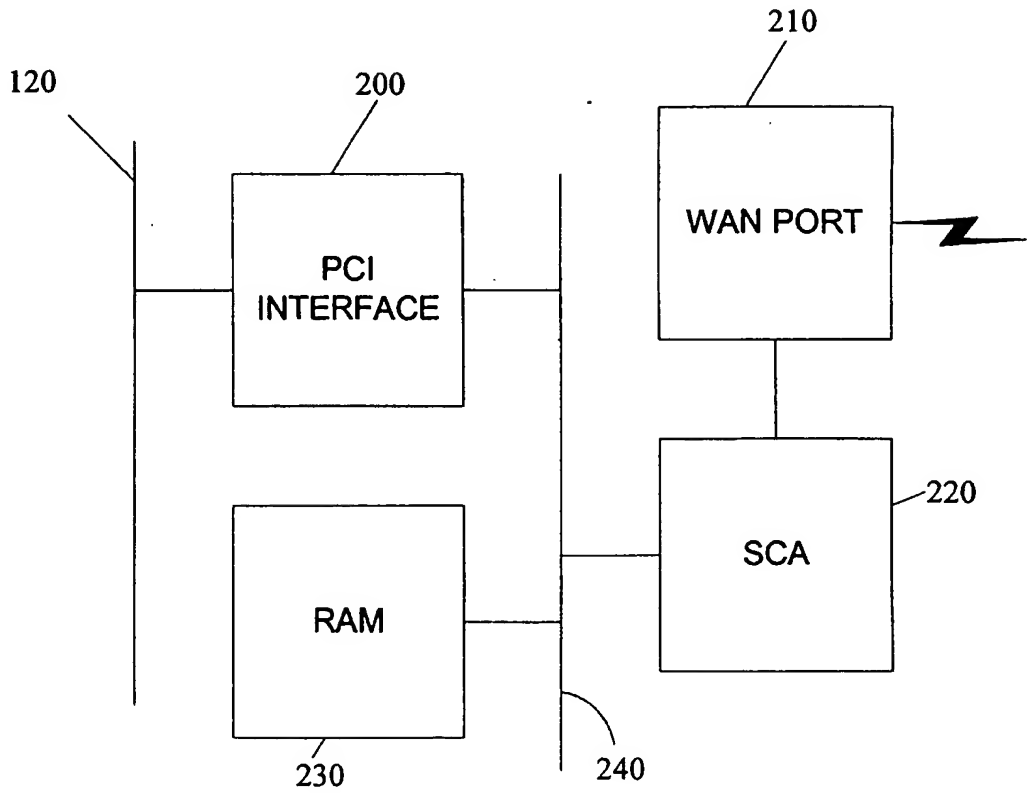
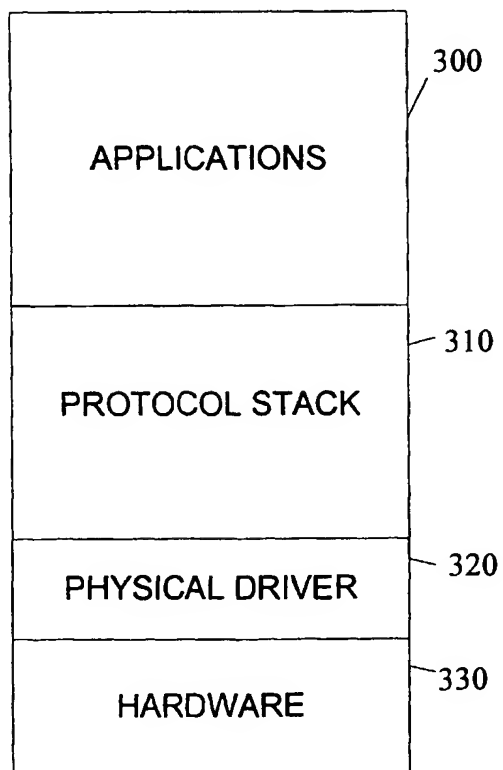


FIG.2

EP 0 903 670 A1

FIG. 3

EP 0 903 670 A1

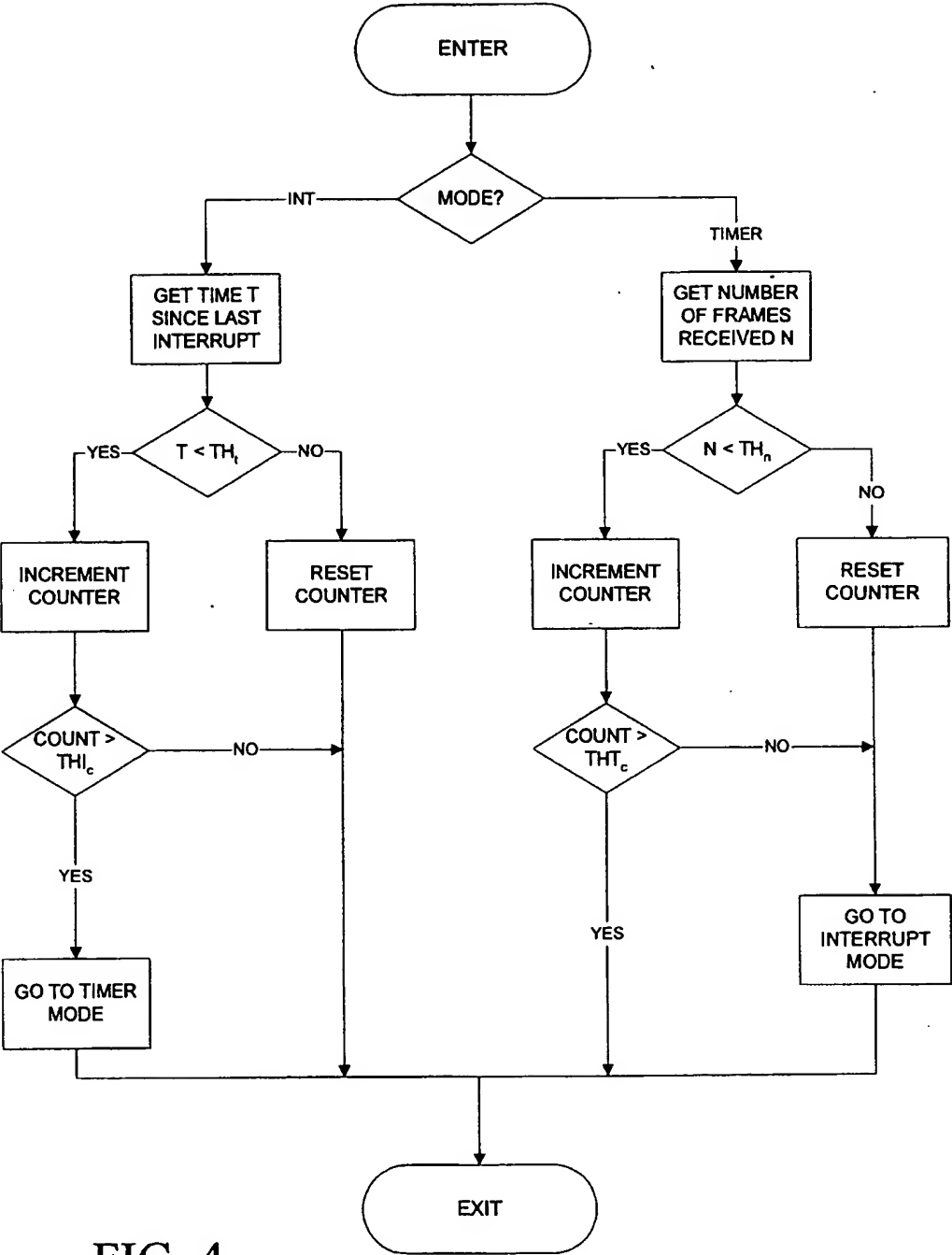


FIG. 4

EP 0 903 670 A1

European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 97 41 0104

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A,D	US 5 414 858 A (HOFFMAN ET AL) * column 3, line 30 - column 4, line 34 * * figures 2-4 * ----	1-9	G06F13/22 G06F13/24
A	US 5 555 414 A (HOUGH ET AL) * column 6, line 49 - column 8, line 13 * * claims 1,6; figures 4A,4B,4C,5 * ----	1-9	
A	"Adaptive Polling Algorithm for Monitoring Multimedia Devices" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 36, no. 9a, September 1993, NEW YORK, US, pages 83-84, XP000395317 * the whole document * ----	1-9	
A	US 5 471 618 A (ISFELD) * column 2, line 1 - column 3, line 23 * * figure 2 * -----	1,8,9	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			G06F
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 12 March 1998	Examiner McDonagh, F
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

EPO FORM 1503 03/82 (P04C01)

EP 0 903 670 A1

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 97 41 0104

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

12-03-1998

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5414858 A	09-05-95	CA 2105020 A	12-06-94
		JP 2520568 B	31-07-96
		JP 7146797 A	06-06-95
US 5555414 A	10-09-96	NONE	
US 5471618 A	28-11-95	NONE	

EPO FORM P0458

For more details about this annex : see Official Journal of the European Patent Office. No. 12/82